Ground-Truth Database for Regional Seismic Identification Research

Lori Grant
Claudia Carabajal*
Multimax, Inc., 331-C Forest Ave., Suite 3, Pacific Grove, CA, 93950

Contract No. F19628-95-C-0094
Sponsored by DOE

Abstract

The Ground-Truth Database (GTDB) provides a subset of the events necessary for seismic discrimination research: accurately located events with regional waveforms and known source type. While it is true that digital waveform data are easily exchanged through the Internet, the data are significantly more useful to discrimination research when event windows have been extracted from the continuous data and when additional documented information, such as source type, is associated with each event. Such a database, readily available and in a consistent format, minimizes the extensive efforts of data collection, organization, review and verification that are often the first phase of new research projects. Having this type of information is essential for development and assessment of algorithms for seismic event location and classification without having to assume the characteristics of the very events under investigation.

The Ground-Truth Database was introduced at the 14th Annual Seismic Research Symposium (Grant and Coyne, 1992) under an ARPA contract. By the time of the 15th Annual Seismic Research Symposium, 288 known events comprising 10 datasets in Western Europe and Scandinavia were available to the research community (Grant et al., 1993b). A large part of the success of the CSS GTDB was due to researchers who wanted to make data available to others by providing the results of their own data-collection efforts to the GTDB.

Two general methods were used to build these datasets. If the target events were the subject of a research study where ground-truth was published for a list of events, then we collected the associated waveforms. Another approach was to look at the regional seismicity bulletins produced by the IMS (Bache et al., 1990) and seek macroseismic information for significant clusters of events. In resuming the compilation of the GTDB we shall follow the model of the previous work but focus mainly on the Middle East and North Africa.

Keywords

Ground-Truth, Regional Waveforms, Discrimination, Database

*now at: State University of New York, Binghamton, NY 13902

Objective

The objective of this work is to compile a standardized database of regional waveforms from known events in a variety of tectonic regions and from a variety of seismic source types. We seek to bridge the gap between raw data and seismic events useful for regional seismic discrimination research by providing:

- event parameters— accurate event location and origin time, phase identifications and arrival times
- digital waveforms— preferably recorded at regional distance
- source type— known with a high degree of confidence, documented

Preliminary Research Results

Previous Effort

This work is a continuation of the Ground-Truth Database previously compiled under ARPA funding at the Center for Seismic Studies (CSS - now known as CMR- Center for Monitoring Research). Over a period of a year and a half, macroseismic information was gathered for waveforms residing in the IMS (Bache et al., 1990) and GSETT-2 (Bratt, 1992) databases.

The CSS GTDB comprises 10 distinct datasets totalling 288 known events as summarized in Table 1 and plotted in Figure 1. With the exception of Dataset 6, the events are grouped into clusters of between 10 and 31 events. Most of the waveforms are from the IMS Arrays ARCES, NORES, FINES and GERES. Some GSETT-2 data are also available for events occuring during that experiment (Bratt, 1992). Various methods were used in collecting ground-truth information: ground-truth data gathered as a part of a published study (Datasets 1,2,5,7,8,10); subsets of other databases (6); ground-truth acquired directly from in-country observers (3 and 4); and datasets where ground-truth information and waveform data were contributed together (9).

The CSS GTDB was released in two versions. Version 1, including datasets 1-3, is described in (Grant et al., 1993a). Version 2, including datasets 4-10, was described in a series of postings and e-mails distributed over the Internet. The data are no-longer on-line at CMR but are availble in CSS 3.0 format upon request to the authors. These datasets have been used in numerous CTBT-related research projects (see, for example the studies of (Pulli and Dysart, 1994; Baumgardt and Der, 1993; Fisk et al., 1993)).

New Effort

Our new data collection effort focuses on the Middle East and North Africa. Figure 2 gives an indication of the natural seismicity in the area of interest (10 to 50 degrees latitude and -30 to 75 degrees longitude). We have plotted all events with magnitude 4 or greater from the USGS database between January 1990 and June 1995. The great majority of these events are related to the tectonically active areas of the Mediterranean Sea, the Hindu Kush and

the Zagros fold belt. When we have access to seismological bulletins with accurate locations for smaller events, the seismicity patterns will presumably include events with other source types such as mining-related events.

The first priority for assembling waveform data from the area would be the high-frequency arrays which are planned in Egypt and Pakistan. As these stations are not yet operating, we will begin the work with other available data. Major networks that may record regional data in the area of interest are shown in Figure 3.

Future Plans

We are in the beginning stages of this project. The first step is to target events that have a chance of meeting the three criteria listed above and gather waveforms for those events. The next step is to verify the source type of the events.

After several datasets have been compiled, they will be carefully analysed for correct phase picks and consistency of event parameters. Following the model of the previous work, we shall work on several different datasets simultaneously. We plan to work closely with other contractors to provide the most useful datasets. One anticipated source of information is the World Wide Web page published at Cornell (Barazangi et al., 1995). Completed datasets will be announced on DOE's Web page, the URL for which is not yet available.

Acknowledgments

Maps were made using GMT (Wessel and Smith, 1991). We apologize for using outdated German and Czech borders.

References

Atakan, K., Lindholm, C. D., and Havskov, J. (1993). Earthquake swarm in Steigen, northern Norway: an unusual example of intraplate seismicity. *Terra Nova.*

Bache, T. C., Bratt, S. R., Wang, J., Fung, R. M., Kobryn, C., and Given, J. (1990). The Intelligent Monitoring System. *Bull. Seismol. Soc. Am.*, 80(6):1833–1851.

Barazangi, M., Seber, D., Vallve, M., Fielding, E., and Isacks, B. (1995). A geological and geophysical information system for Eurasia, the Middle East and North Africa: digital database development for the Middle East and North Africa. PL-TR-94-2092, Phillips Laboratory Scientific Report. ADA297018

Baumgardt, D. R. and Der, Z. (1993). Investigation of regional seismic discriminants using visualization and statistical analysis methods in the Intelligent Seismic Event Identification System. In Lewkowicz, J. F. and McPhetres, J. M., editors, *Proceedings of the 15th Annual Seismic Research Symposium, Vail, Colorado, 8-10 September, 1993*, PL-TR-93-2160, Phillips Laboratory, Hanscom AFB, MA. ADA271458

Bratt, S. R. (1992). GSETT-2: an experiment in rapid exchange and interpretation of seismic data. *EOS Trans. Amer. Geophys. U.*, 73(48).

Fisk, M. D., Gray, H. L., and McCartor, G. D. (1993). Applications of a robust statistical framework for seismic event identification. In Lewkowicz, J. F. and McPhetres, J. M., editors, *Proceedings of the 15th Annual Seismic Research Symposium, Vail, Colorado, 8-10 September, 1993*, PL-TR-93-2160, Phillips Laboratory, Hanscom AFB, MA. ADA271458

Gestermann, N., Harjes, H.-P., Jost, M., Schweitzer, J., and Wüster, J. (1992). GERESS: Monitoring natural and artificial seismicity in Central Europe. In Lewkowicz, J. F. and McPhetres, J. M., editors, *Proceedings of the 14th Annual Seismic Research Symposium, Tuscon, Arizona, 16-18 September, 1992*, PL-TR-92-2210, Phillips Laboratory, Hanscom AFB, MA. ADA256711

Grant, L. and Coyne, J. (1992). Ground-truth data for seismic discrimination research. In Lewkowicz, J. F. and McPhetres, J. M., editors, *Proceedings of the 14th Annual Seismic Research Symposium, Tuscon, Arizona, 16-18 September, 1992*, PL-TR-92-2210, Phillips Laboratory, Hanscom AFB, MA. ADA256711

Grant, L., Coyne, J., and Ryall, F. (1993a). CSS Ground-Truth Database: Version 1 Handbook. C93-05, Center for Seismic Studies Technical Report, Arlington, Virginia.

Grant, L., Ryall, F., and Coyne, J. (1993b). CSS Ground-Truth Database: Update and Case Study. In Lewkowicz, J. F. and McPhetres, J. M., editors, *Proceedings of the 15th Annual Seismic Research Symposium, Vail, Colorado, 8-10 September, 1993*, PL-TR-93-2160, Phillips Laboratory, Hanscom AFB, MA. ADA271458

Joswig, M. and Schulte-Theis, H. (1993). Master-event correlations of weak local earthquakes by dynamic waveform matching. *Geophysical Journal International*, 113.

Kradolfer, U. (1992). Recent seismicity in Switzerland as located by the Swiss network and the GERESS array. In *Proceedings of the GERESS Symposium, Waldkirchen, Germany, 22-24 June.*

Kremnetskaya, E. O. and Trjapitsin, V. M. (1992). Induced seismicity in the Khibiny Massif (Kola Peninsula). 1-92/93, NORSAR Scientific Report, Kjeller, Norway.

Kvaerna, T. (1993). Intelligent post-processing of seismic events— Part2. 2-92/93, NORSAR Scientific Report, Kjeller, Norway.

Mykkeltveit, S. (1993). Mining explosions in the Khibiny Massif (Kola Peninsula of Russia) recorded at the Apatity three-component station. 1-92/93, NORSAR Scientific Report, Kjeller, Norway.

Polish Academy of Sciences (1993). S. J. Gibowicz, personal communication.

Pulli, J. J. and Dysart, P. S. (1994). Two-dimensional signal processing for regional seismic event identification. PL-TR-94-2235, Phillips Laboratory Final Report. ADA292601

Wessel, P. and Smith, W. (1991). Free software helps display data. *EOS Trans. Amer. Geophys. U.*, 72:441,445–446.

Wüster, J. (1993). Discrimination of chemical explosions and earthquakes in Central Europe–a case study. *Bull. Seismol. Soc. Am.*, 83:1184–1212.

Table 1: Ten Datasets of the CSS GTDB

Dataset Name	Quake	Quarry Blast	Mine Tremor	Method	Publication
Version 1					
1: Vogtland	11	15	—	V*	(Wüster, 1993)
2: Steigen	25	—	—	V%	(Atakan et al., 1993)
3: Lubin	—	—	31	V#	(Polish Academy of Sciences, 1993)
Version 2					
4: Silesia	—	—	31	V#	(Polish Academy of Sciences, 1993)
5: Swiss Swarms	27	—	—	V*	(Kradolfer, 1992)
6: Gestermann	10	2	2	D	(Gestermann et al., 1992)
7: Ruhr Basin	—	—	13	*	(Joswig and Schulte-Theis, 1993)
8: Apa. Tremors	—	—	10	*	(Kremnetskaya and Trjapitsin, 1992)
9: Apa. Blasts	—	53	—	*	(Mykkeltveit, 1993)
10: Apa. Blasts	—	58	—	*	(Kvaerna, 1993)
Totals	73	127	88		

V	contact with original in-country source
%	locations from analysis of temp. seismic network
#	locations from mining seismic network
D	subset of another database
*	list of events from published report with ground-truth

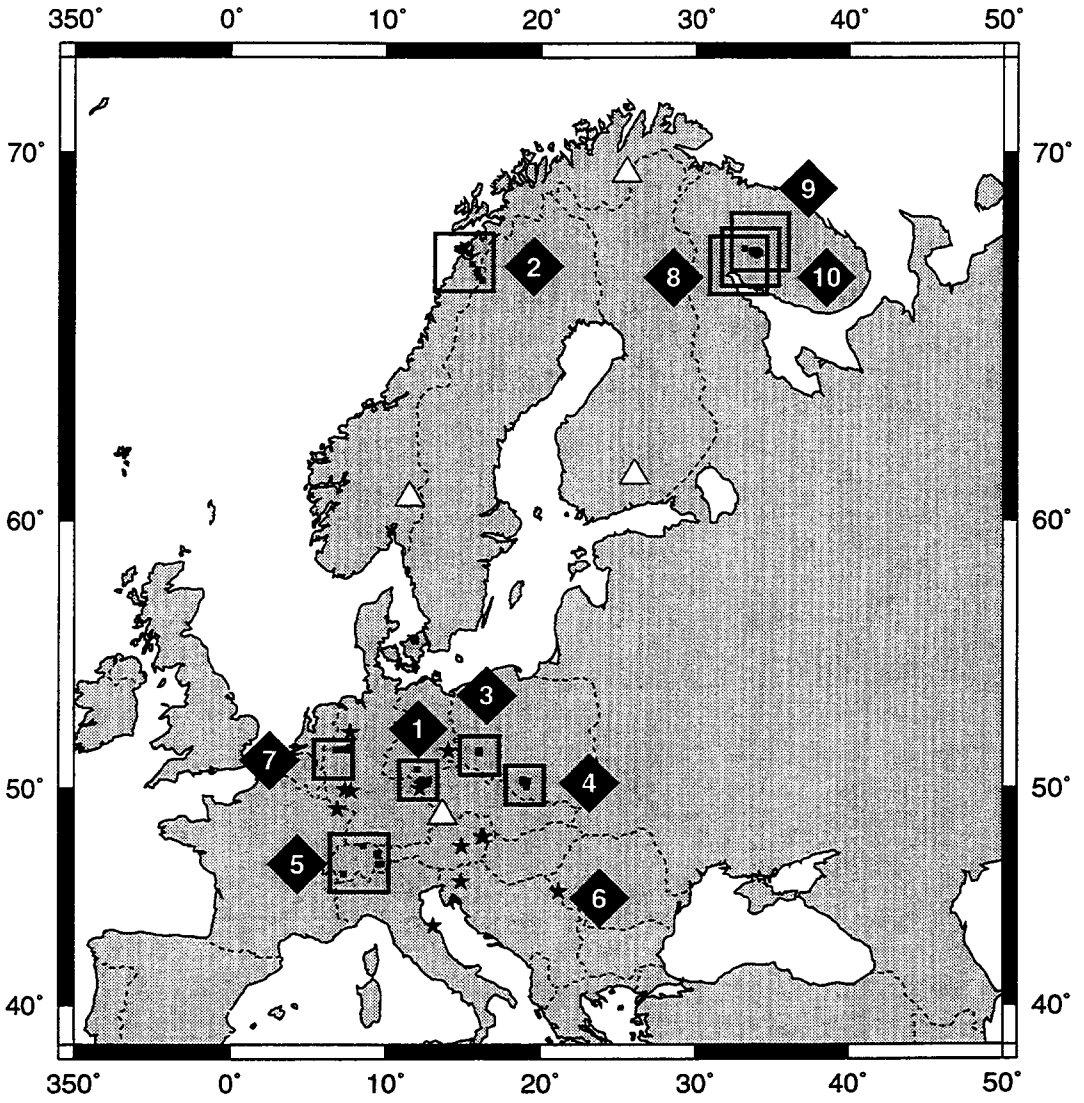

Figure 1. Ten Datasets of the CSS Ground-Truth Database. Open squares enclose clusters of events known as datasets 1-5 and 7-10 which are numbered by diamonds. Dataset 6 event locations are shown by stars. The majority of waveforms for these events comes from the high-frequency arrays ARCES, NORES, FINES and GERES, indicated by white triangles. Some waveform data from GSETT-2 stations are also available. See Table 1 for summary and text for details. This database is available from the authors (lori@es2.multimax.com).

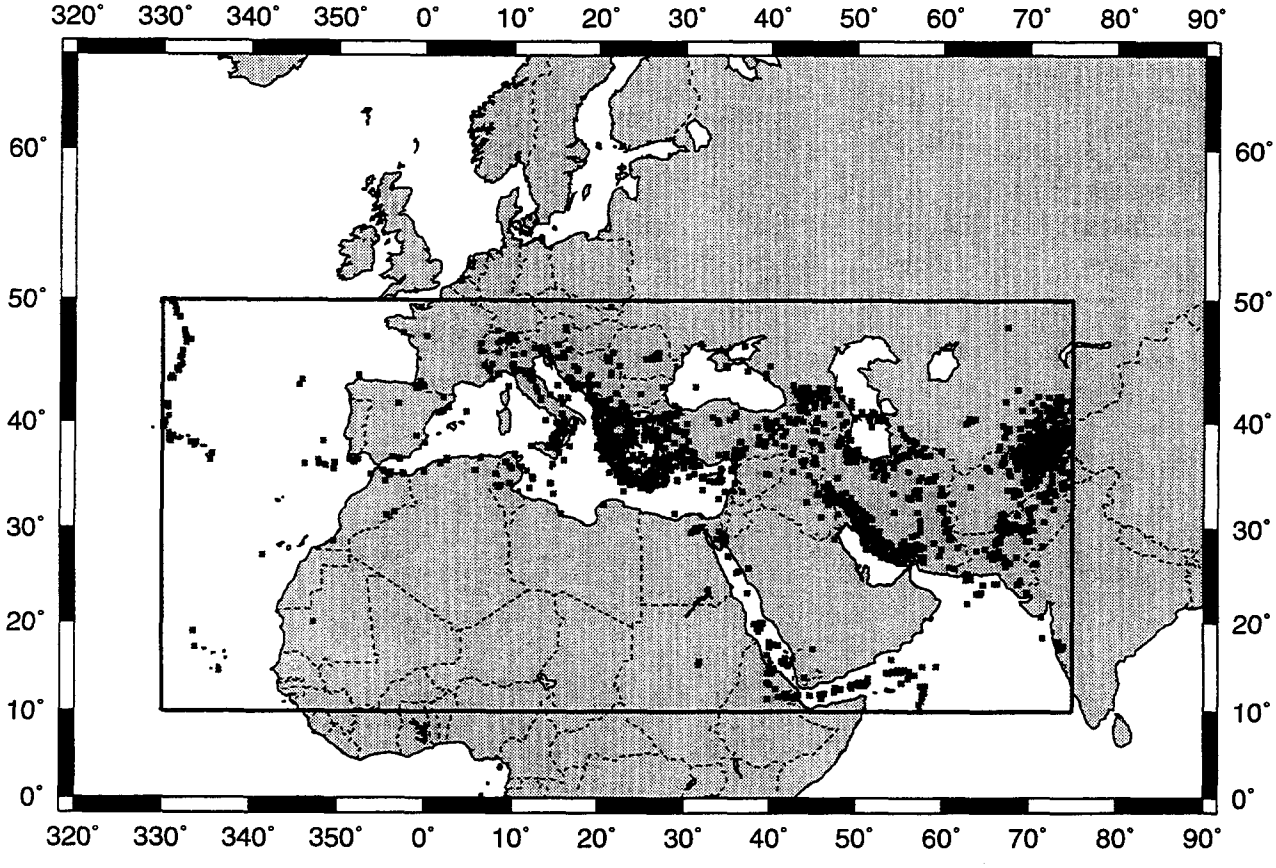

Figure 2. Seismicity in area of interest. USGS locations for all events with magnitude greater than 4 between January 1990 and May 1995.

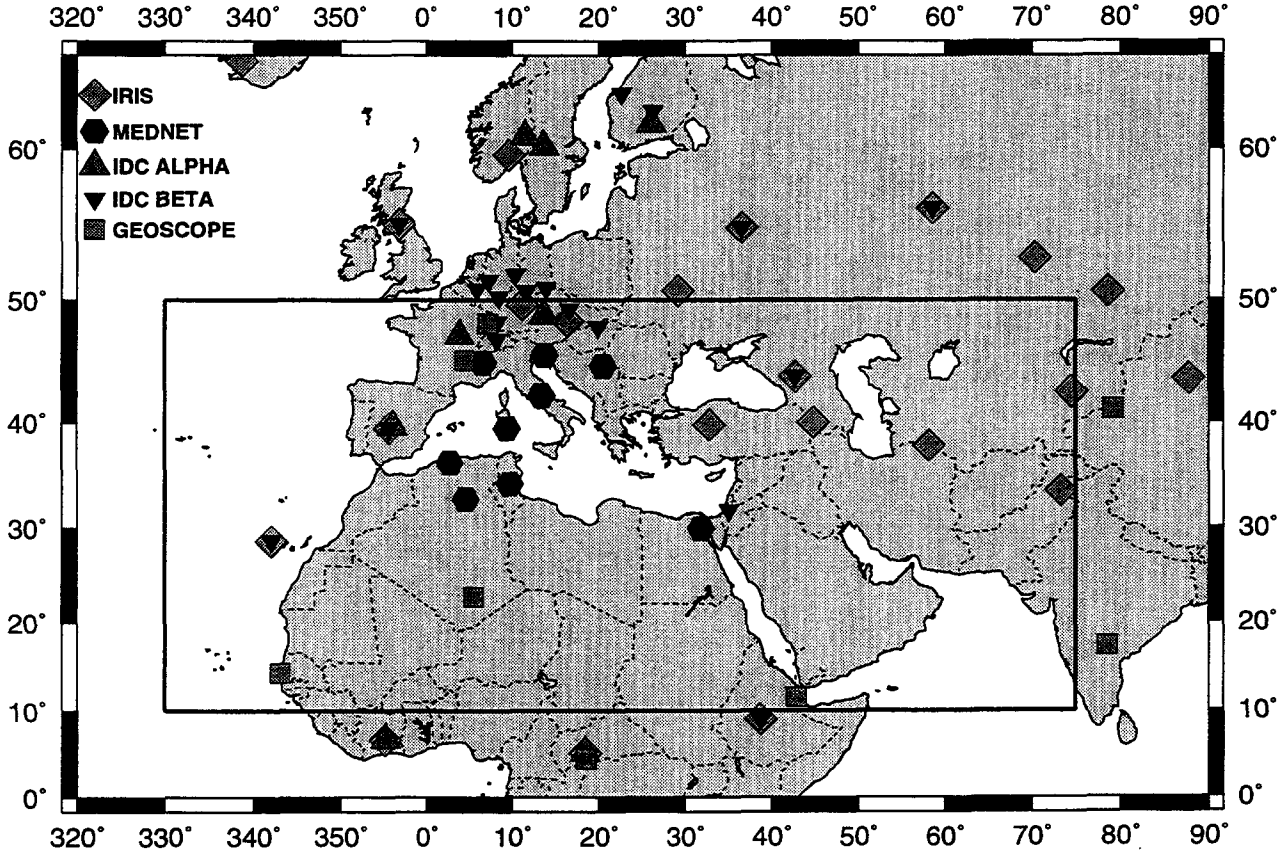

Figure 3. Networks that may record regional data in the area of interest. The IRIS network includes IRIS/USGS and IRIS/IDA stations. Many of the stations are affiliated with more than one of the networks listed. This is indicated on the map by overlapping symbols. An array is planned in Egypt appoximately 4 degrees southeast of the Mednet station KEG. The Pakistan array is planned at the same location as the current IRIS/IDA station NIL.